5-8-2015

Wrong Plus Wrong Equals Right: Exploring the Interaction between Bacteriophage T7 Exonuclease and ssDNA-Binding Protein Using Suppressor Mutations

Stephanie Delzell

Wrong Plus Wrong Equals Right: Exploring the Interaction Between Bacteriophage T7 Exonuclease and ssDNA-Binding Protein Using Suppressor Mutations

Stephanie Delzell

Submitted in Partial Completion of the
Requirements for Departmental Honors in Biology

Bridgewater State University

May 8, 2015

Dr. Boriana Marintcheva, Thesis Director
Dr. Michael Carson, Committee Member
Dr. Meredith Krevosky, Committee Member
Dr. Jenna Mendell, Committee Member

BRIDGEWATER STATE UNIVERSITY

Wrong Plus Wrong Equals Right: Exploring the Interaction Between Bacteriophage T7 Exonuclease and ssDNA-Binding Protein Using Suppressor Mutations

---

Stephanie Delzell
Mentor: Boriana Marintcheva

Table of Contents

Introduction

A Brief Overview of Viruses

Viruses are non-living biological entities with enormous diversity that are able to infect organisms from every kingdom of life. Viral infection in humans is the cause of many well-known diseases, such as influenza and herpes. These viruses are notorious for their detrimental effect on individuals and society. However, many viruses also serve a wide variety of ecological roles, and can serve to benefit us in the biotechnology and medical fields if we know how to utilize them properly.

Viruses are evolutionarily prepared to infect single cells of an organism and utilize the cellular machinery to replicate. Viruses do not contain organelles of their own, and rely on host cells for all of their protein production. The composition of a virus includes a protein capsid encasing a nucleic acid genome, which may be ssDNA, dsDNA, ssRNA, or dsRNA. Some viruses also have a phospholipid envelope. In order to hijack the cell, viruses have developed different strategies of entry, some of which involve the entry of a whole virus into the cell and others only injection of viral nucleic acid. Viruses are then able to use various specialized molecular mechanisms to replicate their own genome and produce necessary viral proteins in the cell. New viral progeny is assembled, and released from the host cell, killing it.

Bacteriophages

Bacteriophages, or phages, are viruses that use bacterial cells as their hosts. They were discovered independently by two individuals, the first of whom was Frederick W. Twort. Twort recognized that a filterable agent was causing the death of bacterial cells on plates. He did not use the term “virus” to characterize his observations, although at the

time filterable infectious agents of plants and animals were known. He wrote a paper in 1915 that is now regarded as the birth of modern bacteriophage research, but it was disregarded for about five years after it was published. Frederick d'Herelle received credit for the discovery of phages, after describing the lysis of bacteria in liquid medium in 1917. D'Herelle was interested in the potential use of this bacteria-killing agent to fight disease caused by bacteria, and his work following the discovery of phages was focused on their biological properties, and their use as therapeutic agents. The pursuit of phages as antimicrobials largely died out with the development of antibiotics in the 1930s.

Bacteriophages are the most common biological entities on the planet, with an estimated population size of $1x10^{31}$ individuals, outnumbering bacteria ten to one. (Shors et al., 2013). Of all these phages, over ninety-six percent of them are tailed phages with dsDNA belonging to the order Caudovirales, which use their tails as conduits for genetic material being injected into their bacterial host (Ackerman et al., 2011). Caudovirales can be separated into three families; *Myoviridae*, which have long, contractile tails, and includes bacteriophage T4 and T4-like phages, among others. The second family under the umbrella of Caudovirales is *Podoviridae*, which are characterized by short, non-contractile tails, and include T7 and T7-like phages. The third family is *Siphoviridae,* including bacteriophage lambda, which have long, non-contractile tails, and are known for their distinct genetic code. Other phages are not tailed, falling outside these three orders, and may be filamentous or polyhedral in morphology (Ackerman et al., 2011). Bacteriophages, as such a large and diverse group, are hard to

classify definitively. However, important features such as morphology and genetic makeup can contribute to our understanding of their relationships to other bacteriophages.

Two types of lifecycles, lytic and lysogenic, have been described in the world of viruses. Both involve specific viral molecular strategies for hijacking host cellular machinery to support viral replication. Lytic phages immediately begin creating progeny and kill host cells rapidly. Temperate or lysogenic phages may exist in a latent state, incorporating their genome into the hosts, and existing there temporarily as a prophage. These cells are free to live and divide, indirectly replicating the viral genome as they do. Eventually, latent viruses reactivate and begin propagation effectively switching to a lytic life cycle. Both types of lifecycles begin with a stage called adsorption, in which the virus attaches to the host via receptors on the cell surface. Phages often accomplish this using their tail fiber proteins to recognize these receptors. Next the viral genome is injected into the host cell, and in the case of tailed phages, the genetic material travels down the tubular structure into the cell. In lytic phages, biosynthesis would begin quickly, and the phage utilizes the host's resources to replicate its own genome and translate new proteins. The new phages are assembled in the cell, and then are released back into the environment.

Bacteriophages in general have a limited host range, known to successfully infect few host species. This property allows phages to be used in effectively in identification of bacterial serotypes, or variations within bacterial species, which are often hard to distinguish using conventional methods. The ability of a phage to infect and lyse certain bacteria assists researchers in narrowing down the possible identities of an unknown organism (Haq et al., 2012).

Phages in the Environment

Bacteriophages are common in the environment, and can be found in many of the expected locations of their hosts, and commonly in greater numbers. In oceans and freshwater lakes, phages outnumber bacteria in a volume of water, and also have a greater diversity (Chibani et al., 2004).

Marine bacteriophages are recognized as a critical component in bacterial carbon recycling through the lysis of bacterial cells. Obtaining more information about how these phages work will only contribute to our understanding of global ecology. *Synechococcus* and *Prochlorococcus* are two examples of Cyanobacteria that are infected by a type of phage known as Cyanophages. These two bacterial genera alone are responsible for about half of the world's primary production, making the phages that infect them of interest at a large scale to determine their impact on global carbon cycling (Clokie et al., 2011).

Bacteria are commonly found in the guts of animals, including mammals, and often aid in digestion and contribute to an organism's overall health. The ecosystem found within an animal's gut is commonly referred to as its gut microflora. Phages are also known to be part of this gut microflora. However, quantitation of the bacteria in the guts of these animals demonstrates higher concentrations of the prokaryotes than the presence of phages normally allows. It has been suggested that the GI tract environment might be prohibitive for phage adsorption, due to factors such as pH, nutrient availability, relation to mucosa, and the possibility of the presences of phage inhibiting substances. Bacteriophages are best studied in the microflora of non-ruminant mammals, and there is

still a lot to be learned about their role in the gut ecology of other types of animals, such as reptiles and amphibians (Letarove et al., 2009).

As with many known predator-pair relationships in nature, bacteriophages are often considered a driving force in the evolution of their bacterial hosts. Phages can drive bacterial evolution much as another predator might, by selecting for individuals that are able to evade infection by phages (Breitburt et al., 2005) In addition, phages have the ability to transfer genes from one bacterium to another, contributing to the genetic diversity of a population of bacteria (Breitburt et al., 2005). Some phage genomes code for virulence factors that can be transferred to bacterial genomes, and can change the bacterial phenotype to a pathogenic one. For example, a pathogenic strain of *E.coli,* O104:H4, gains its pathogenicity in the form of a gene that codes for Shiga toxin, which is transferred from a phage genome (Turner et al., 2011). This is one reason careful consideration is required when selecting phages to use as treatments or components of a treatment in humans.

Bacteriophages in Biotechnology

Due, in part, to their molecular simplicity, phages can be manipulated for directed use in various applications of modern biotechnology. Applications of phages discussed herein include phages as anti-microbials, phage display, and utilization of phage-coded enzymes for use in laboratory applications.

Phages as Anti-Microbials (Bacteriophage Therapy and Biocontrol)

Upon first recognition of bacteriophages as infectious agents of bacteria, it is a natural inclination to attempt to use them for the lysis of pathogenic bacteria. Indeed, these applications can be successful and were a focus of research when phages were first

recognized for their potential in the 1920s, but was largely abandoned when conclusions about the efficacy of such treatments were variable, and antibiotics were developed in the 1930s (Shors et al., 2013). In modern medicine, antibiotic resistant strains of pathogenic bacteria have begun cropping up, making new modes of attack necessary. This has turned attention back to the potential of bacteriophages to combat bacteria.

One concern this strategy raises is whether it can be expected that infectious bacteria will develop resistance to phage therapy, since bacteria have been shown to be able to rapidly develop resistance to phages in laboratories. This occurs most commonly via alterations of the bacterial extracellular receptors (Örmälä et al., 2013). The development of resistance may not be as problematic as antibiotic resistance since bacteriophages are a diverse and dynamic group, capable of change that could alter their ability to infect bacteria, just like they do in nature.

Bacteriophages being used as an anti-microbial tool may have applications in treating human diseases, but also show promise in food-safety and agricultural applications. For instance *Xanthomonas pruni* is a bacteria known to cause spots on peaches, and has been successfully controlled by phage-mediated biocontrol (Haq et al., 2012). Phages have also been used to control Listeria contaminations in meats. Because phages are so commonly present in the environment, these treatments are not required to be listed by food manufacturers, as it is presumed phages would already be present in significant quantities in the products in question (Hudson et al., 2009). It is likely that these techniques will become more refined over the coming years, and possibly more universally adopted.

Phage Display

Phage display is a technique with a variety of applications. When a peptide sequence is attached to a phage's major capsid protein using molecular biology strategies, hundreds of copies of the engineered protein can be displayed by a single phage. Phage display techniques have been used to create combinational libraries that contain high quantities of different phage clones displaying various encoded peptides on their capsids (Samoylova et al., 2009). These phages can be sorted through to find polypeptides that will have high affinity for target proteins through biopanning. The technique is commonly used to identify contact surfaces between proteins and to select for high affinity antibodies.

Another unique application of phage display is to label phages for studying the processes and behavior of phages themselves. Having T4 phage display a fluorescent protein (GFP) aided in the visualization of the phage and allows the study of its interactions with living mammalian tissues and organs in the context of phage therapy. (Kaźmierczak et al., 2014).

Phage Enzymes Commonly Used Nucleic Acid Manipulation

Enzymes that are commercially marketed for laboratory use are often products of phage genomes. These enzymes can have properties not found in other forms of life that make them more useful in specific applications and protocols. These enzymes are often surprising in their versatility, as some can be used outside of the expected prokaryotic applications.

DNA polymerases are often utilized in molecular biology techniques that require the replication of DNA. One polymerase that is used from phages is T7 DNA

polymerase. The protein acts in two ways, both as a polymerase, and a 3'-5' exonuclease (Grippo et al., 1971). T7 DNA polymerase is ideal for making copies of long stretches of DNA. The commercially available version of this enzyme is created using two components, one is gp5 of T7 bacteriophage, the other is Thioredoxin from *E. Coli.* Genetically engineered T7 DNA polymerase is known as sequenase and is used in dideoxy sequencing. Another phage-originating DNA polymerase used is T4 DNA polymerase. It has the same two functions as the T7 version, and is regarded as the best polymerase for use in creating blunt ends (Tabor et al., 1989).

When ligation is necessary in laboratory protocols, several phage-coded ligases are commercially available, all with different specific properties. Ligase is used for forming bonds in the sugar-phosphate backbone of nucleic acids (Figure 1). T4 DNA ligase can be used to join blunt-end termini of DNA or RNA by catalyzing the formation of a phosphodiester bond of a 5' end and a 3' end that are located next to each other. It can also repair nicks in the backbone of double stranded molecules of DNA, RNA, or DNA/RNA hybrids (Engler et al., 1982). Another phage ligase available is T3 DNA ligase. T3 DNA ligase works with dsDNA. It serves a similar function to that of T4, but it exhibits a higher tolerance for NaCl in a reaction (Cai, 2004). T7 ligase is also commercially available for ligation of dsDNA, but it does not catalyze blunt-end ligation efficiently. T4 RNA ligases are also available, and can be used for ligation of ssDNA or ssRNA. Examples of applications for using these enzymes are making single-stranded oligonucleotides or incorporating unnatural DNA sequences into genes (England et al., 1977). The range of specifications available from phage enzymes is useful for obtaining tools for procedures that have specific requirements.

An example of a DNA degrading enzyme that is commercially available is T7 exonuclease, which is a product of bacteriophage T7 gene 6 (described in detail below). This enzyme is able to degrade dsDNA starting at the 5' end, or starting from nicks or gaps. It can also degrade RNA and DNA in RNA/DNA hybrids, but cannot degrade ssRNA or dsRNA.

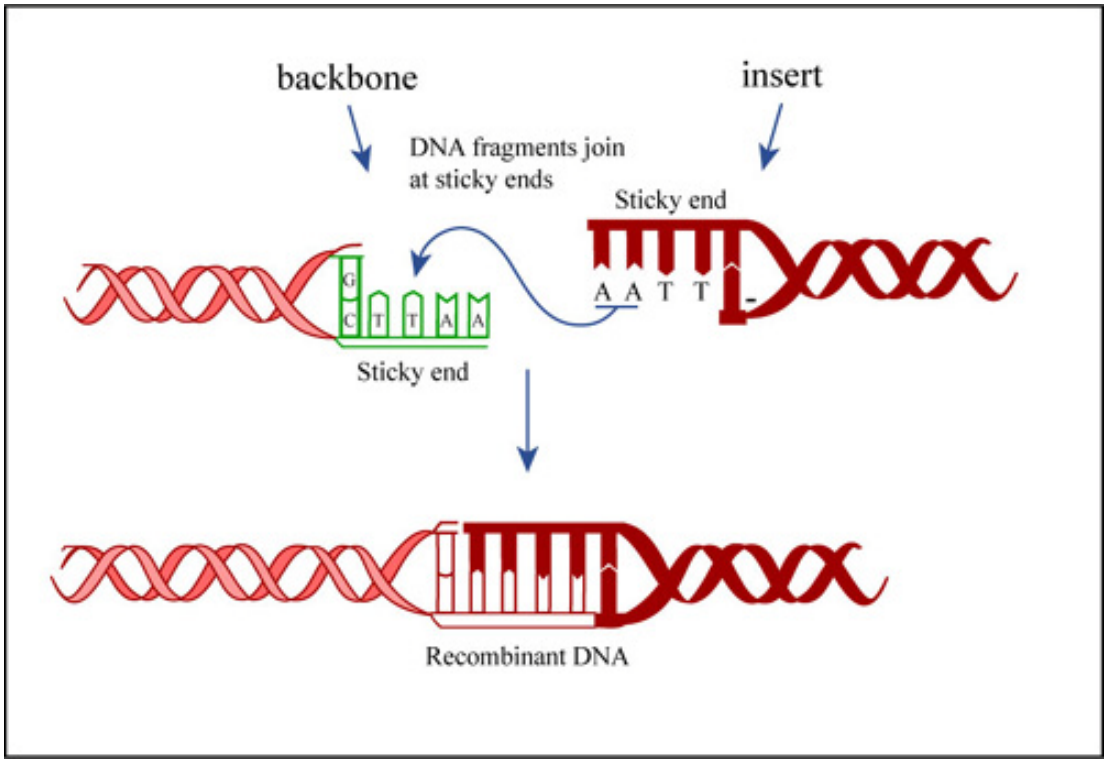

***Figure 1**: **DNA Ligation.** Ligase catalyzes the formation of the bonds in sugar-phosphate backbone of the DNA after the nitrogenous bases have formed Hydrogen bonds with each other. This figure was originally published in Endy et al., 2007.*

CRE Recombinase from bacteriophage P1 is interesting in its ability to function in eukaryotes, making it useful in creating recombinant animals, such as in transgenic mice, to study gene function. CRE Recombinase catalyzes site-specific recombination at a 34 bp sequence called loxP. The result of CRE Recombinase on two strands of DNA that each have a loxP site would be the fusing of these two DNA strands (Abremski et al., 1984). Uses of phages in biotechnology are still in development, but specific strategies here have been in the works for years or decades, and have demonstrated a degree of success. Knowing that there is so much to be learned in the world of phages, it is logical to expect that as science develops a clearer picture of phage biology that more applications will emerge for use in the field.

T7 Bacteriophage

Bacteriophage T7 is a member of the order Caudovirales, which consists of tailed bacteriophages, and the family Podoviridae, which is characterized by a short, non-contractile tail and a linear, double-stranded DNA genome (Ackerman et al., 2011) (Figure 2). This phage infects *E.coli* in a lytic cycle, and is among the relatively few well-studied phages in biology. This is partially due to the wide availability of its host and the ability to culture it easily in a laboratory setting (Clokie et al., 2011). The tails of Podoviridae bacteriophages vary from species to species, but have similarities that seem to suggest a common mechanism for DNA delivery into the host cell. These similarities include a tubular structure surrounded by fibers. The T7 tail is composed entirely from the arrangement of four proteins, including connector protein (gene product 8, or gp8), gatekeeper protein (gp11), nozzle protein (gp12), and a fiber protein (gp17) (Cuervo et.

al., 2011). Investigating these proteins can demonstrate T7's relationship to other viruses, and help establish functions for the individual proteins involved in T7, and the homologous proteins found on other Podovirales phages. This phage, as one with a wealth of information already known about it, provides the opportunity to ask much more specific questions about the functions and mechanisms of its proteins. T7 has a complex viral shape, and a dsDNA linear genome. The entire genome of T7 bacteriophage is approximately 40,000 base pairs in length and codes for 55 proteins.

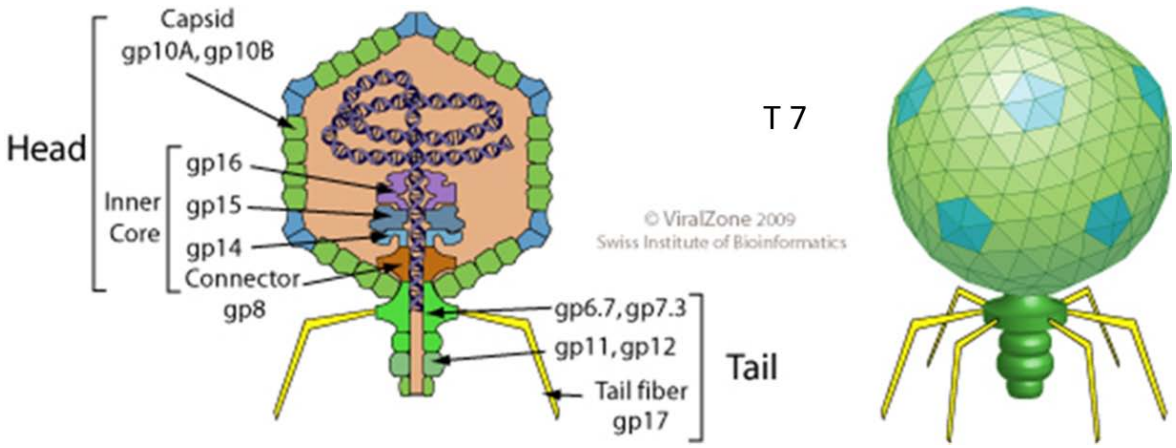

Figure 2: Illustration of a T7-Like Bacteriophage. *Gene products that make up the tail, inner core, and capsid are labeled and color-coded, as well as a whole outside illustrated view. This figure was originally published by the Swiss Institute of Bioinformatics, 2009.*

T7 Replisome and GP6

All of the proteins that work together to replicate DNA in one synchronized machine are collectively known as a replisome. Although some of the T7 replisome proteins are utilized by biotechnology, not all of the characteristics and interactions of the specific proteins are fully understood. *In vitro,* only four T7 proteins are necessary to

replicate DNA. These proteins are gp5 (polymerase), gp2.5 (ssDNA binding protein), *E. coli* encoded thioredoxin, and gp4, which functions as both a helicase and a primase, (reviewed in Lee, et al., 2011 and Hamdan et al., 2009, Figure 3). The only protein in the T7 replisome that is not encoded by the phage's genome is the thioredoxin (trx), which serves as a processivity factor for T7 polymerase, enabling it to stay associated with primer-template for the addition of many more nucleotides before dissociation than gp5 by itself (Hamdan et al., 2009).

Gp2.5, or single stranded DNA binding protein, is an instrumental protein for the replication process. It has a structure that allows it to coordinate replication proteins by exposing its C-terminal tail when it is bound to ssDNA (Figure 4, Marintcheva et al., 2008). This ability to coordinate proteins makes gp2.5 a key player in replication, recombination, and DNA repair, and raises the question of how it is able to interact with the other proteins.

Phage Encoded		**Host Encoded**	
Gene	**Function**	**Gene**	**Function**
Gp 5	DNA polymerase	Trx A	Processivity Factor of Gp 5
Gp 4	Helicase/Primase		
Gp 2.5	ssDNA Binding Protein		

Figure 3: Necessary Proteins for T7 DNA Replication In Vitro.

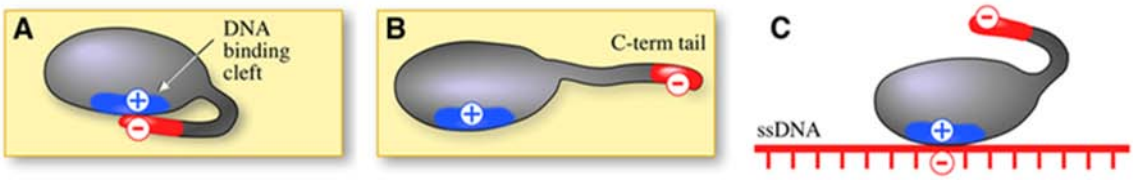

Figure 4: GP2.5 Function. *The C-terminal tail of ssDNA binding protein is able to bind to the positively charged cleft where ssDNA can bind. When the protein is bound to DNA, the tail is freed to interact with other proteins. This figure was originally published in Marintcheva, et al., 2007.*

Because of the directionality of DNA synthesis, (DNA is synthesized 5' to 3'), replicating both anti-parallel strands of the double helix simultaneously results in a lagging strand and a leading strand. The leading strand is able to be continually replicated by polymerase, while the lagging strand must be replicated in short pieces as the template is exposed by the helicase. These short pieces formed on the lagging strand are called Okazaki fragments. Single-stranded DNA binding protein is essential for stabilizing the lagging strand of DNA during replication, and in the T7 replisome, it has also been implicated in organizing the other players in replication to be able to simultaneously replicate both strands of DNA with the same machine. The lagging strand of DNA is looped in order to come in contact with the replisome machinery (Figure 5). Gp2.5 is known to directly interact with gp5 (DNA polymerase), and gp4 (helicase/primase) (reviewed in Marintcheva et al., 2006). The specific interactions within the replisome are not well understood. One approach used to learn more about the protein interactions, and reveal physical sites of interaction, is a suppressor mutation screen.

Marintcheva et al. performed a suppressor mutation screen using mutations in the C-terminal tail of gp2.5. The suppressor mutation screen was performed by introducing gp2.5 lethal mutations into the phage genome and selecting for a viable phage. T7

carrying lethal gp2.5 mutation can become viable only if the mutation is suppressed or reversed. These new compensatory mutations are called suppressor mutations, as they suppress the lethal phenotype (Figure 6). Gp2.5 is an ideal candidate for performing such a screen because it interacts with many other proteins to coordinate their functions. One of these GP2.5 mutations that produced extragenic mutations swapped the positions of amino acids aspartate, which is polar and hydrophilic, and phenylalanine, which is non-polar and hydrophobic. Two point mutations in gp6, W42C and E120D, were shown to suppress the lethal phenotype mutation with both mutations present on the gene (Marintcheva et al., 2009).

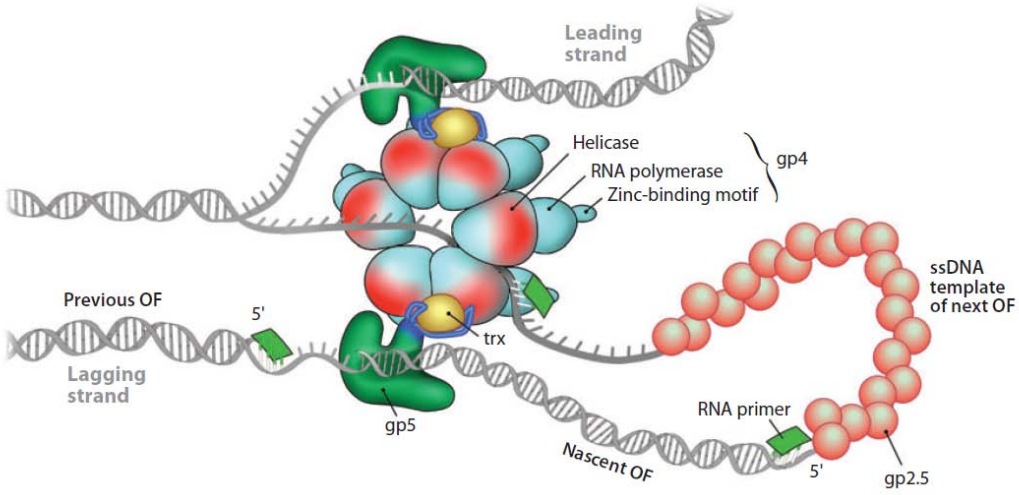

Figure 5: The T7 Replisome. *Gp2.5 is seen in pink, and plays a role in coordinating the lagging strand with the replisome by stabilizing the ssDNA loop. Gp5, polymerase, is shown in green, and helicase in red and blue. Figure originally published in Hamdan, et al., 2009.*

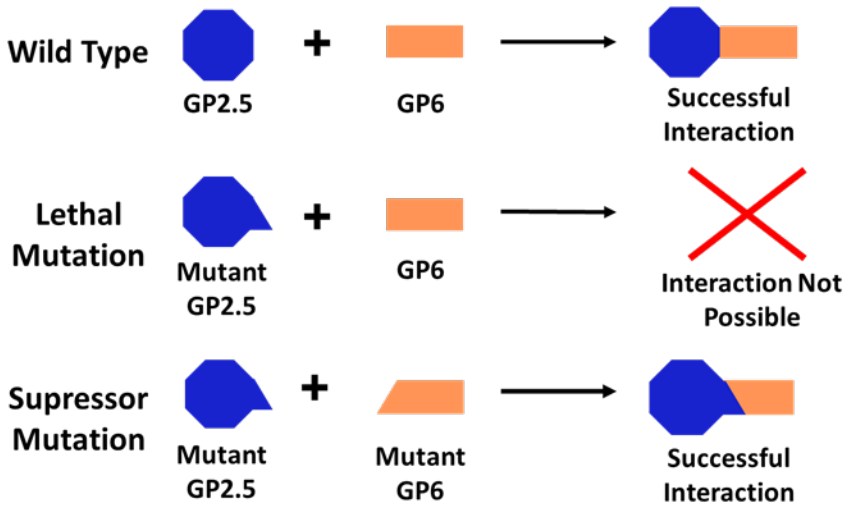

Figure 6: Suppressor Mutation Molecular Mechanism. *The suppressor mutation screen revealed potential interactions between GP 2.5 and GP6, when two mutations in GP6 acted to suppress the lethal phenotype created by the GP2.5 mutation, so the phage was viable, as it is with wild-type proteins.*

T7 exonuclease, or gp6 of bacteriophage T7, is a commercially available DNA degrading enzyme. This enzyme is able to degrade dsDNA starting at the 5' end, or starting from nicks or gaps. It can also degrade RNA and DNA in RNA/DNA hybrids, but cannot degrade ssRNA or dsRNA. The enzyme, when functioning as an exonuclease, enables the breakdown of the host genome for use in the assembly of the T7 genome. Gp6 can also directly aid in rescuing stalled replication forks, enabling replication to continue (Figure 7).

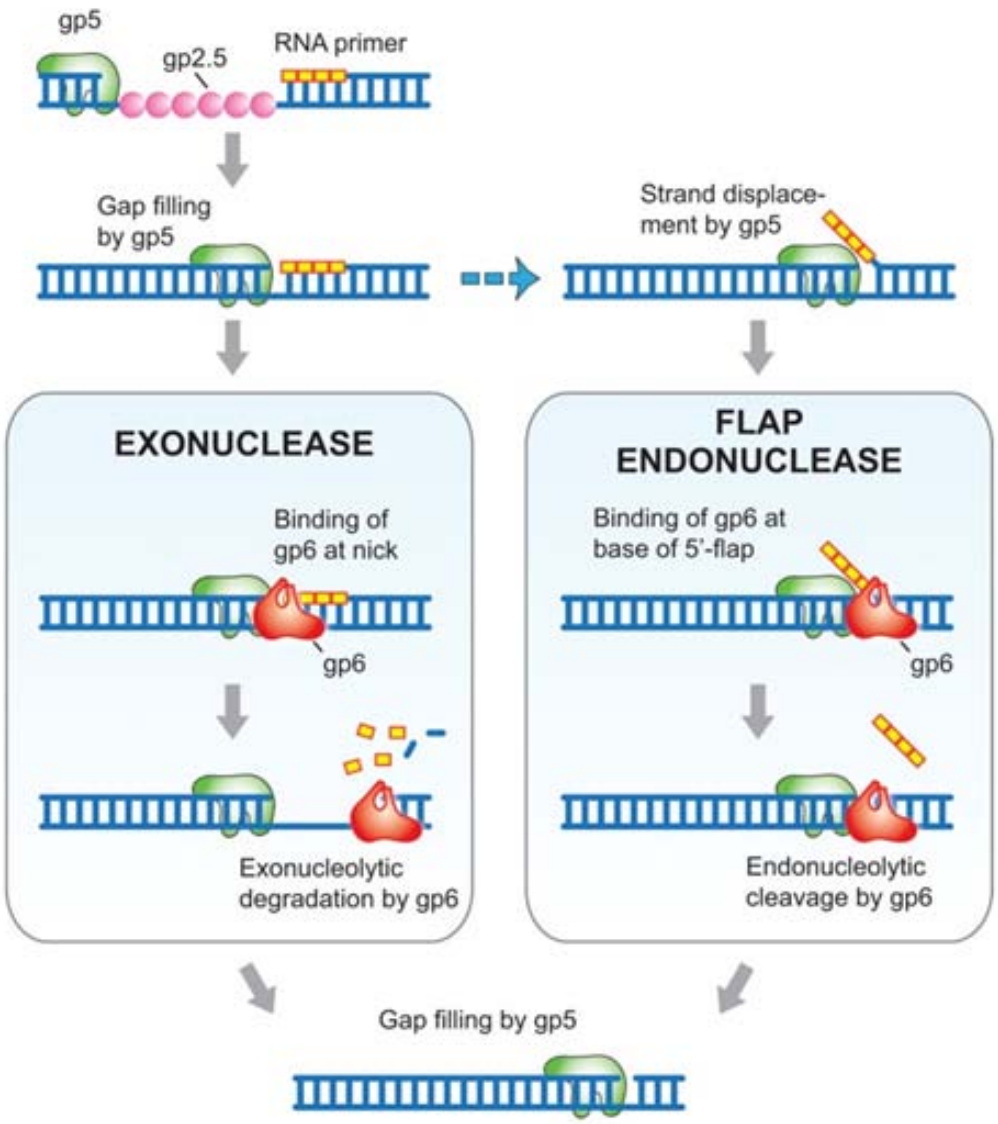

Figure 7: Gp6 Functions. *Gp6 is depicted in red, gp5 in green and gp2.5 in pink. The figure displays two functions of GP6, as an exonuclease and a flap endonuclease. This figure was originally published in Mitsunobu et al., 2015.*

Research Goal

The two point mutations in gp6 described above (W42C and E120D) are what the research herein is poised to investigate, by re-creating these mutations using PCR strategies. In the first mutation, tryptophan (W) is replaced with cysteine (C) at the 42nd position in the amino acid sequence. Tryptophan is a bulky and hydrophobic amino acid,

while cysteine is more compact, and hydrophilic. This mutation would therefore change how the protein could interact with others, and likely the causative agent of the suppressor phenotype. In the second mutation, glutamic acid (E) is replaced with aspartic acid (D) at the 120th amino acid location. Both of these amino acids are hydrophilic and charged, while glutamic acid has a slightly longer side chain. These amino acids would interact with other proteins in a similar manner, and therefore this mutation is less likely the causative mutation of the suppressor phenotype.

In order to determine definitively whether one of the two mutations or both are necessary for the suppressor phenotype, three versions of the gene must first be generated; a version with mutation W42C, a version with E120D, and a version with both mutations. These mutated versions of gene 6 will ultimately be inserted into plasmids for later use in complementation and suppressor assays. These assays will be used to determine whether both gene 6 mutations must be present to suppress the lethal phenotype induced by the mutation in gp2.5, or if either W42C or E120D can cause the suppressor phenotype on its own.

These interactions are important to understand thoroughly because of the proteins' applications in biotechnology, and also because the T7/*E. coli* relationship is a well-studied phage host relationship, and can lay a foundation for information about how phage replication proteins work. The groundwork laid by this research will ultimately lead to a better understanding of how gp2.5 works, specifically in how it is capable of coordinating activity at the replisome by interacting with gp6.

Materials and Methods

Generation and Evaluation of a New T7 Laboratory Stock

Soft Agar Plating

In soft agar plating, 0.5 ml bacterial culture and 3 ml melted soft agar and viral stock are mixed and then poured on a plate with a hard agar base. Once poured, the soft agar solidifies, allowing bacterial growth in a thin monolayer. Viral plaques are seen as clear areas free of bacteria in a field that is otherwise milky with bacterial growth. When multiple dilutions of viral stock are to be plated on the same plate, the viral dilutions are dispensed in labeled locations on top of the solidified agar, instead of being part of the soft agar mix.

Generation of a new T7 Phage Stock

Wild type T7 bacteriophage was a generous gift from Dr. Charles Richardson, Harvard Medical School. In order to generate a new viral stock, it was first necessary to perform plaque purification to ensure the final stock would be uniform, as started from a single virus. Plaque purification was achieved by picking a single plaque from the soft agar plate with the tip of a micropipette and rinsing the tip in media. The resulting mixture was used to begin a serial dilution to 10^{-5}, and the dilutions were each plated. Several dilutions were plated to ensure the isolation of single plaques. After three rounds of purification, a single plaque was selected to be used in the generation of a new viral stock.

A culture of *E. coli* strain C600 in LB media was infected with T7 phage to form a new viral stock. One single plaque from the final round of plaque purification was stabbed with a micropipette tip and rinsed in LB media. This resulting mixture was used

to infect an *E. coli* C600 culture that had light absorbance at 600nm between 0.1 and 0.2. The infected culture was incubated at 37°C without agitation for half an hour so that the phages would be able to attach to the bacterial cells and inject their genetic material. Next, the culture was incubated at 37°C with shaking. Measurements of absorbance were taken every half hour of the infected culture and the corresponding control bacterial culture (treated exactly the same, but containing no virus) until the infected culture was completely lysed as evidenced by absorbance readings close to zero.

Determining Phage Titer

Determination of the number of phages in the resulting viral stock was accomplished by doing serial dilutions of the stock. The dilutions were all plated on a separate section of a single plate and allowed to grow in the dark overnight at room temperature. They were incubated in the dark in order to prevent unnecessary DNA damage from exposure to UV light. After incubation, the plate was evaluated to determine which dilution would be most likely to form single plaques when plated individually, and in the evaluation of the original stock (Figure 8). Dilutions that were determined likely to yield individual colonies were spread over an entire plate, to find a plate with a number of plaques that fell in the statistically usable range of 20-200 plaques. The dilution used was 10^{-5}, so to determine the number of phages in a volume of stock, the number must first be multiplied by 10^5. The number of plaques counted would be used to determine the phage concentration using the following formula:

$$Viruses/mL = \frac{Number\ of\ Plaques\ x\ 10^5 \left(\frac{plaques}{\mu L}\right)\ x\ 10^3 (\frac{\mu L}{mL})}{Volume\ Stock\ Plated\ (\mu L)}$$

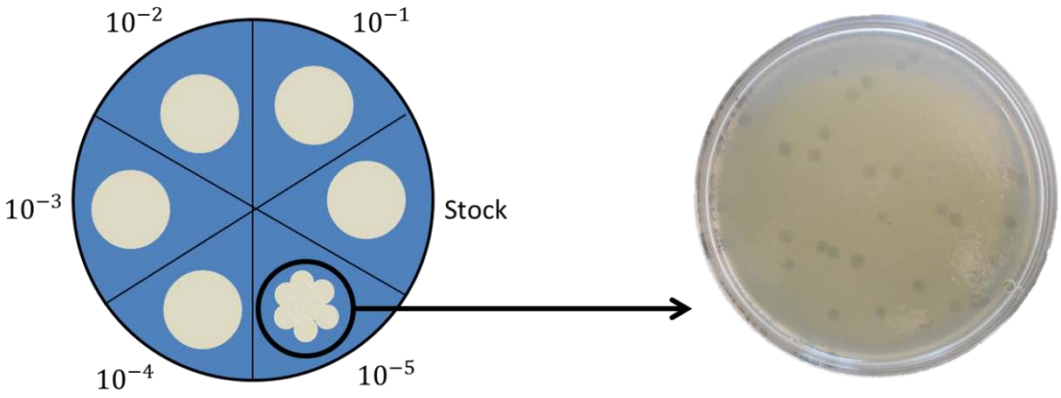

Figure 8: Quantification of the Generated Viral Stock. *Multiple dilutions on the same plate (left) were utilized to determine which dilutions would yield single plaques if spread over an entire plate (right). These plaques could then be counted to determine the viral titer.*

PCR Mutagenesis

Primer Design

Primers were designed for use in both introducing the specific point mutations in GP6 as well as replicating the whole gene via PCR (Figure 9). Pairs of primers were designed to introduce each of the two mutations. The mutagenic primers contained the desired new sequence flanked by the base pairs complementing the wild type sequence, so during PCR amplification the primers would introduce a different base pair where they bound. The forward primer at the start of the gene begins with the sequence "CACC" so that the resulting gene would contain the matching base pairs to be inserted into the pENTR/D-TOPO plasmid later.

In PCR I, oligos in each reaction were used to amplify one of two parts of the gene, as well as introduce the point mutation. For example, one of the PCR I reactions that generated the tryptophan to cysteine mutation used the forward primer binding to the beginning of the gene, and a reverse primer containing the desired mutation surrounded

by the appropriate sequence surrounding the 129th base pair where the mutation is located. The second reaction used the forward primer for the location of the mutation, and the reverse oligo for the end of the gene. This strategy results in two pieces of the gene that have overlapping sequences where the desired mutation is positioned.

A PCR program was designed that accounted for the melting point of the primers used. The PCR program included the following steps: an initial, one-time melting temperature of 95°C for 5 minutes, and then a repeated cycling of a denature step at 95°C for 15 seconds, an annealing temperature of 61.2°C for 30 seconds, and an extend step of 68°C for one minute. After this cycle repeated thirty times, the samples were stored at 4°C.The first PCR reactions (0.3 µM oligos 1 and 2, 1 µL template DNA, 1X reaction mix buffer, 2.5 units/50 µL polymerase, and 40.5 µL distilled water) served to create the mutations, using a total of four reactions, each of which had a total volume of 50 µL. The second PCR reactions (0.3 µM oligos 1 and 2, 10 µL each templates 1 and 2, 1X reaction mix buffer, 5 units/50µL polymerase, and 21µL distilled water) replicated the entire gene, one reaction for each of the two mutations.

```
Start Primer, Forward, Binds to nucleotides 1-38 of Gene 6
Oligo Sequence: 5′CACCATGGCACTTCTTGACCTTAAACAATTCTATGAGTTACG
Gene Sequence:  5′    ATGGCACTTCTTGACCTTAAACAATTCTATGAGTTACG

W→C Oligo, Forward, Binds to nucleotides 122-144
Oligo Sequence: 5′  GAGTTTGATGCCTCTTGCGAGGAAGAGATTTG
Gene Sequence:  5′  GAGTTTGATGCCTCTTGGGAGGAAGAGATTTG

W→C Oligo, Reverse, Binds to nucleotides 122-144
Oligo Sequence: 5′ CCAAATCTCTTCCTCGCAAGAGGCATCAAACTCAGC
Gene Sequence:  3′ GGTTTAGAGAAGGAGGGTTCTCCGTAGTTTGAGTCG

E→D Oligo, Forward, Binds to nucleotides 346-378
Oligo Sequence: 5′ CTTGATGCTCTTTGACCGCGAAGAGTTCTATTGC
Gene Sequence:  5′ CTTGATGCTCTTTGAGCGCGAAGAGTTCTATTGC

E→D Oligo, Reverse, Binds to nucleotides 346-378
Oligo Sequence: 5′ ATAGAACTCTTCGCGGTCAAAGAGAGCATCAAGG
Gene Sequence:  3′ TATCTTGAGAAGCGCGAGTTTCTCTCGTAGTTCC

Stop Primer, Reverse, binds to nucleotides 879-903
Oligo Sequence: 5′ CTACGGTCTCCACAGGTAAATCTCC
Gene Sequence:  3′ GATGCCAGAGGTGTCCATTTAGAGG
```

Figure 9: Primers Designed to Introduce Each of the Two Point Mutations. *Differences between the original gene and the oligo are highlighted. In the Start primer, the CACC sequence is necessary to base pair with the overhang sequence in the vector. The differences in the mutagenic oligos (W→C and E→D primers) are the point mutation being introduced through the replication of the sequence with these oligos.*

Gel Electrophoresis

One percent agarose gels were used to separate DNA fragments, both to evaluate the success of the PCR reactions based on the resulting fragment sizes, and extract the DNA fragments to be purified and used. Gels were run using 1X TBE buffer at 100 Volts, for roughly 45 minutes each, unless otherwise noted. DNA was visualized in the gels with the use of 0.001mg/mL ethidium bromide in the agarose gel and a UV light box.

Gel Purification

QIAquick extraction kit (Qiagen) was used to extract the PCR fragments from the agarose gels once they were determined to be of the correct length. Bands were cut from the gel and then the buffers of the kit were used to dissolve and extract the gel according to the manufacturer's recommendations.

TOPO Cloning

Topoisomerase was utilized in TOPO cloning reactions that would insert gene 6 with the CACC overhang into the desired vector (Figure 10). Two cloning reactions using different amounts of PCR product were performed for each mutation, in order to insert the genes into pENTR/D-TOPO vector (Invitrogen). All four reactions were allowed to incubate at room temperature for 30 minutes.

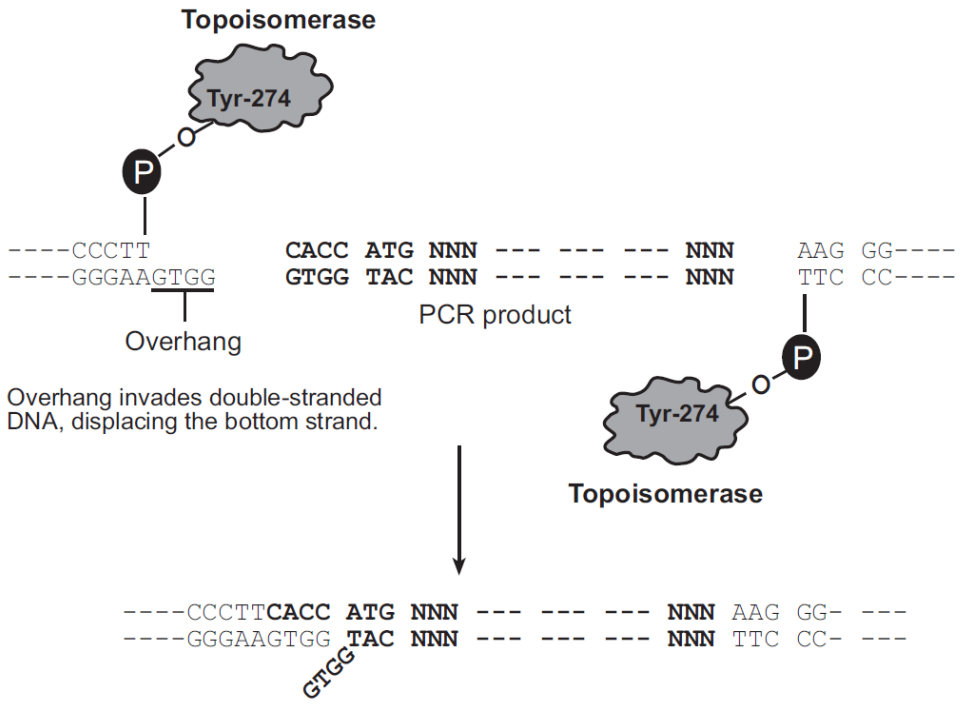

Figure 10: The Mechanism of Topoisomerase in the TOPO Cloning Reaction. *Topoisomerase is used to bind the created PCR product into the pENTR/D-TOPO vector, which could be subsequently transformed into competent cells. Picture originally published by Invitrogen.*

Transformation

Cloning reactions were transformed into OneShot Top 10 chemically competent cells (Invitrogen) by incubating the cells with the cloning reaction product on ice for 15 minutes, heat shocking the cells for 30 seconds in a 42°C water bath, and immediately returning to ice. 250 µL S.O.C. medium was added to each vial of cells, and these were incubated with shaking at 37°C for one hour. The resulting cultures were plated in two different volumes for each culture on kanamycin containing media. The kanamycin ensured only cells that contained the plasmid would grow. Once colonies grew, they were used directly to inoculate new bacterial culture.

Bacterial Culture Growth

A solution of five milliliters of LB media with 0.006mg/mL kanamycin was inoculated with a single colony and incubated with shaking overnight at 37°C.

Plasmid Isolation and Gene Verification

A QIAprep Spin Miniprep Kit (Qiagen) was used according to the manufacturer's recommendations to isolate plasmids from the cells. Plasmids isolated from these cultures served as a template for PCR, using the same program as previously described, to amplify the inserted gene. These PCR products were run on a 1% agarose gel to ensure the presence of a DNA fragment of the size we expected. Plasmids with verified insert were sent for sequencing.

Sequence Analysis

Samples of each of each plasmid verified by PCR was sent to Eurofins Genomics to be sequenced using the dideoxy approach. Once sequencing results were received,

BLAST was run to compare the received sequences with the original gene 6 sequence. This allowed the verification of the presence of the expected mutations, and also the intact gene. Forward and reverse primers were used to verify viable copies of the gene.

Results

T7 Viral Stock Generation and Quantification

Three rounds of plaque purification were performed and a single plaque selected to infect bacterial cultures for the purpose of viral stock generation. The infection was monitored with measurements of light absorbance (OD600, directly monitoring the density of the bacterial culture). For the first generation of stock, an initial absorbance of 0.13 was upon initial infection. The measurements of light absorbance gradually increased as bacteria were dividing and reached a plateau when the rates of cellular division and lysis reached equilibrium. As the rate of cell lysis increased due to the progression of the infection, the optical density declined. The uninfected control measured at the same time points showed a steady increase in absorbance as bacteria replicated unimpeded by the virus, resulting in higher readings. Once the absorbance on the infected sample read close to zero, it was assumed all bacteria were lysed and the infection was complete, leaving the phage suspended in media. The second generation of culture was treated the same as the first, with at an initial absorbance of 0.076.

The first infection, or generation I, was performed in a volume of 10mL. This generation was used to infect a 50mL generation 2, using the same absorbance measurements to observe the progress of the infection. To ensure generation 2 would be a viable laboratory stock, the density of this new suspension needed to be determined. To do this, serial dilutions of the stock to 10^{-11} were made and 5µL of each dilution was plated in its own section on one of two plates.

When it was determined a third generation would be necessary to further increase the density of the stock, a higher concentrated bacterial stock was used (about 0.2

absorbance) and the viral stock was diluted to 1:10, its original strength. These alterations were made to try to reduce the rate of viral infection, and allow for more rounds of replication resulting in higher viral count. This generation was incubated, shaken, and monitored as the previous two. The quantitation plates showed that we would likely get single plaques by plating dilutions 10^{-5} and 10^{-6} which were plated separately to determine the density of the new stock.

Once the third generation stock dilution 10^{-5} was plated by itself, it had 37 single plaques, which is in the ideal range for use in calculations. This number was used in the formula in Material and Methods, leading to the final figure of $7.4 * 10^{8}$ viruses/mL. Overall, the produced phage stock was at the lower range for a good viral stock, but with a usable concentration.

PCR Introducing Mutations

Primers were designed to incorporate two different point mutations into GP6 at 129 base pairs (bps) and 363 bps. These mutagenic primers are represented by the green and blue arrows that overlap the mutation site. The primers on the ends of the gene can also be used to later amplify the whole gene (Figure 11). In order to introduce the tryptophan (W) to cysteine (C) mutation, W42C, at the 129th base pair, the gene was split into pieces two pieces via PCR I of sizes 144 base pairs and 792 base pairs. The PCR I fragments to create the glutamate (E) to aspartate (D) mutation, E120D, are 558 and 378 nucleotides in length. These fragment sizes were verified by a gel (Figure 12) PCR II was designed to reassemble gene 6 from the overlapping fragments, to result in products that are the full size of the gene, 903 base pairs. The PCR products from PCR 1 were gel purified and used as templates to reconstitute the entire gene in PCR II amplification. In

order to introduce the second mutation in one plasmid, the same two PCR procedures were used, using the genes created with single mutations as the template.

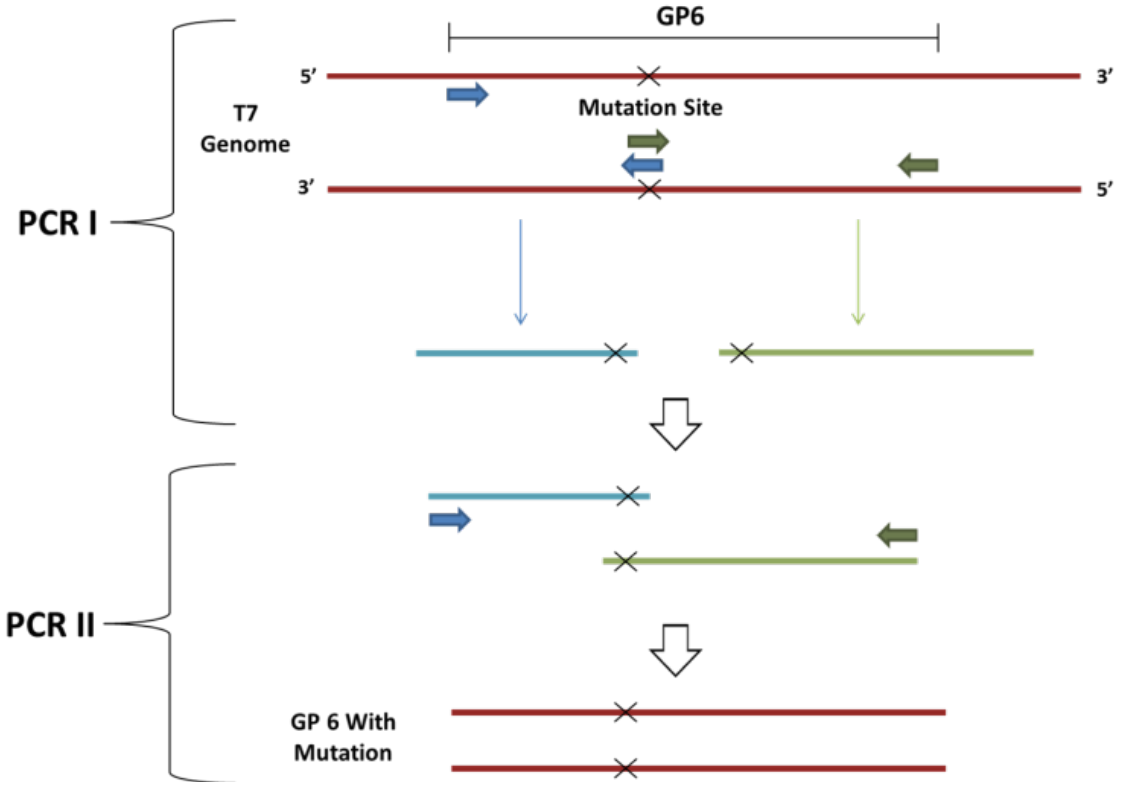

Figure 11: PCR Strategy for Introducing Point Mutations. *In this figure, blue arrows follow the amplification of one gene fragment while green arrows follow the amplification of the other gene fragment. These two pieces each contain the mutation, and when they are reassembled in PCR II, the amplified gene contains the desired mutation.*

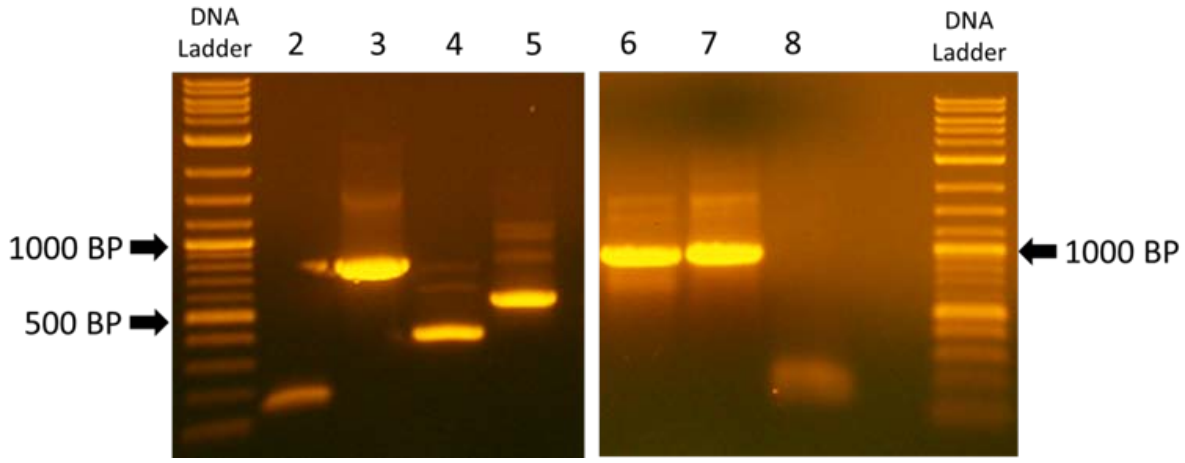

Figure 12: Gel electrophoresis of PCR 1 and 2 Products. *Lanes 2 and 3 of PCR 1 show the W to C mutation products, 144 base pair length in lane 2 and 792 in lane 3. Lanes 4 and 5 in PCR 1 have the second mutation, for E to D. Lane 4 is 558 base pairs and lane 5 is 378. These fragments are reassembled into the whole gene in PCR 2. The gel for PCR 2 shows the successful reassembly of gene 6 in lanes 6 and 7. The gene is 903 base pairs long. Lane 8 is the positive control at 130 base pairs. New England Biolabs' Quick-Load 2-Log DNA Ladder was used.*

Cloning of gp6 Variants Containing Single Mutations

After gene 6 was amplified in PCR with the desired mutations, the gene was inserted into the pENTR/D-TOPO vector (Invitrogen) using an array of concentrations of PCR product so that the optimal fragment to plasmid ratio could be found to yield the best results (Figures 13 and 14). These reactions were transformed in to OneShot Top 10 chemically competent cells (Invitrogen) yielding cells containing the plasmid with the kanamycin resistance gene.

Isolation of Transformed E.coli

After undergoing transformation, the cell cultures were plated on kanamycin-containing media to select for cells that had the plasmid. This step ensured that the cells grown would all contain the plasmid, which contains the kanamycin resistant gene.

Colonies were counted on the plates to be able to verify the number of colonies would be congruent with the volume of culture plated. For all reactions with one mutation, viable colonies grew, meaning all cloning reactions resulted in the uptake of plasmids (Figure 12) These colonies were used to inoculate culture to be used for the isolation of plasmids containing the single mutations. These plasmids were tested for the desired insert via PCR using primers to amplify gene 6 before they were sent for sequencing.

For the reactions that used the gene containing two mutations, overall fewer colonies grew, but some colonies grew for every reaction (Figure 13). The growth of colonies indicated successful cloning as well as successful transformation. These colonies were used to inoculate new cultures from which plasmids containing both mutations would be isolated, used as a template in PCR to verify the gene insert, and sent for sequencing.

	PCR Product	TOPO Vector	Number of Colonies (50μL)	Number of Colonies (200 μL)
Tube 1 (W→C)	0.5 μL	1 μL	32	101
Tube 2 (E→D)	4 μL	1 μL	40	84
Tube 3 (W→C)	1 μL	1 μL	34	101
Tube 4 (E→D)	3 μL	1 μL	22	54

Figure 13: TOPO Cloning Reaction and Resulting Colonies for the Introduction of a Single Mutation. *Each of the four reactions were set up with different quantities of PCR product and water, so that in the case one ratio did not work, an optimal reaction would be easier to find, saving time if troubleshooting were necessary. This table also shows how many colonies were counted on each plate resulting from the spread of transformed E. coli onto selective plates, verifying the successful transformation of pENTR/D-TOPO plasmid into the competent cells.*

Cloning Reaction	PCR Product	TOPO Vector	Number of Colonies (50μL)	Number of Colonies (200 μL)
Tube 1a	0.5 μL	1 μL	9	65
Tube 1b	4 μL	1 μL	22	64
Tube 2a	1 μL	1 μL	5	29
Tube 2b	3 μL	1 μL	0	15
Control	N/A	N/A	0	0

Figure 14: Composition of Reactions and Resulting Colonies from Transformation of Plasmids Containing gene 6 with Two Mutations. *Each plate was counted for colonies that contained the plasmid, as a way to verify the efficacy of the transformation process and evaluate cells for plasmid content and gp6 content (achieved through plasmid isolation and PCR amplification of gene 6 using previously designed primers.) The control used in this experiment showed no growth on the kanamycin plates, but this was due to an error in procedure. The control plasmid was resistant to ampicillin rather than kanamycin, and displayed growth on the appropriate plate.*

PCR Determining the Presence of Gene 6 in Isolated Plasmids

A PCR verification step was utilized to demonstrate the insertion of T7 gene 6 in the pENTR/D-TOPO plasmid. The PCR products for the first round of reactions for which the plasmids each contained the gene with one of two mutations had uniform band patterns which verified the presence of the gene, but also had much larger DNA present which was not identified (Figure 15). These unexpected bands created some concern for the quality of the plasmids, but it was decided to send the plasmids for sequencing. Because the results were uniform, two samples of each mutation were chosen arbitrarily for sequencing.

The PCR verification for the insert of the gene with two mutations was less uniform, with isolations 1, 4, 5, 6, 7, 9 and 10 all with large DNA fragments present. Samples 2, 3, and 8 showed clearly the presence of the gene without other DNA, so these samples were chosen for sequencing. 1, 4, 9, and 10 were also sent for sequencing, because they each had the correct fragment size for gene 6 in the gel. Samples 5, 6, and 7 were not sequenced because all bands present in the gel were too large to be gene 6 (Figure 16).

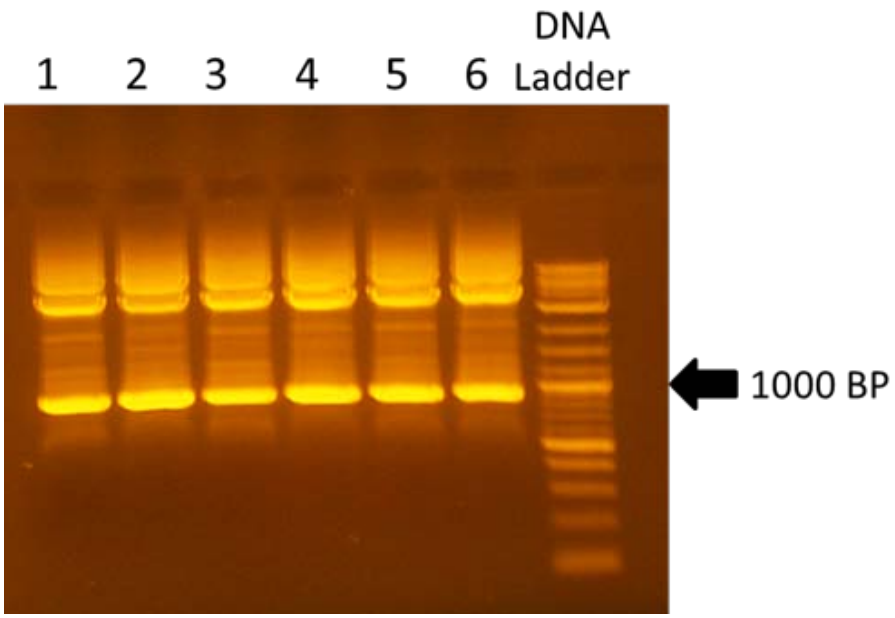

Figure 15: Gel Electrophoresis to Visualize Gene 6 in Isolated Plasmids*. This gel shows the PCR results to amplify gene 6 in the isolated plasmids. Plasmids from all six inoculated cultures were isolated. Lanes 1-3 contain the PCR products from the W to C cultures, and lanes 4-6 contain the PCR products from the E to D cultures. All six show a bright band where we expect the gene to appear, at around 1000 base pairs. New England Biolabs' Quick-Load 2-Log DNA Ladder was used.*

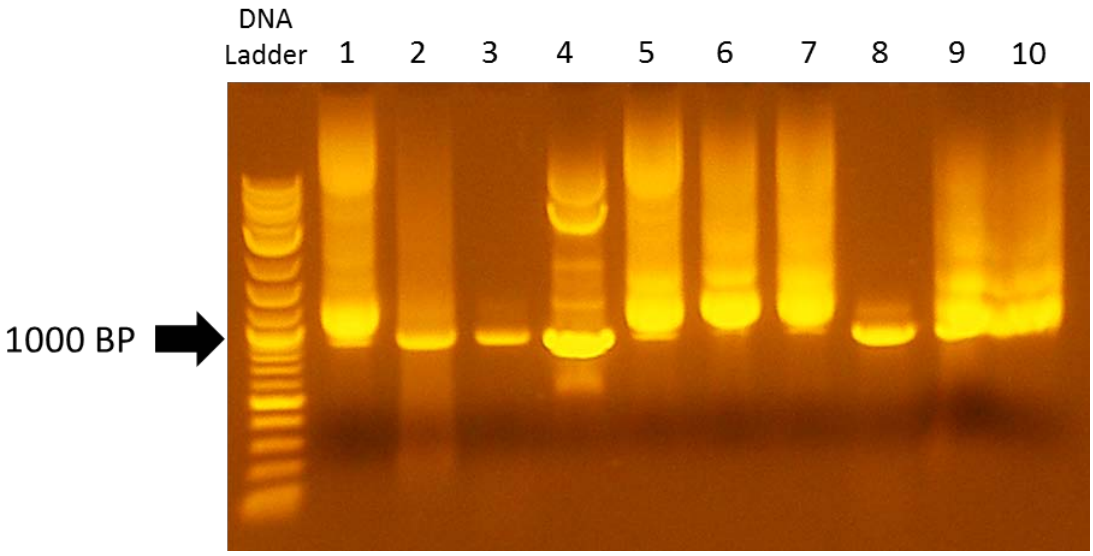

Figure 16: PCR Verification of Gene 6 Insert in Plasmid for Gene Containing Double Mutations. *Lanes 2, 3, and 8 represent samples that were easily selected for sequencing based on the clear results, showing a gene-sized fragment without noise. New England Biolabs' Quick-Load 2-Log DNA Ladder was used.*

Sequence Analysis

The Basic Local Alignment Search Tool, or BLAST, can be used to find similar sequences of DNA, and align them based on similarity. This tool was used to compare the wild type T7 gene 6 with the sequenced gene. In BLAST sequence alignment, lines tying two bases together signify that they are the same. No line appears when the two bases are different. These notations helped find any issues with the sequenced gene, and helped identify the expected mutations where they ought to be. Any issues in the sequenced DNA identified using BLAST could be examined in the chromatogram for the corresponding sequence to determine if there was a problem with the gene.

Analysis of the sequenced DNA using BLAST led to the conclusion that the samples labeled W1 and D1 both had the entire gene intact, included start codons and stop codons, each with one of the two expected mutations Sample W1 contains the W42C mutation (Figure 17). Sample D1contains the expected E120D mutation (Figure 18). Because these two samples each have of one of the two desired point mutations, they would be viable for use in future assays.

Plasmids that were expected to contain both of the mutations were also sequenced, and then used in BLAST. Of the seven sequences received, only those labeled 4 and 10 had enough continuous sequence to see the whole gene. Of these, only 10 had an entire intact gene, with the start codon and stop codon, and also contained both of the expected mutations (Figure 19). This sample will be viable for use in future assays that require a version of the gene that has both mutations.

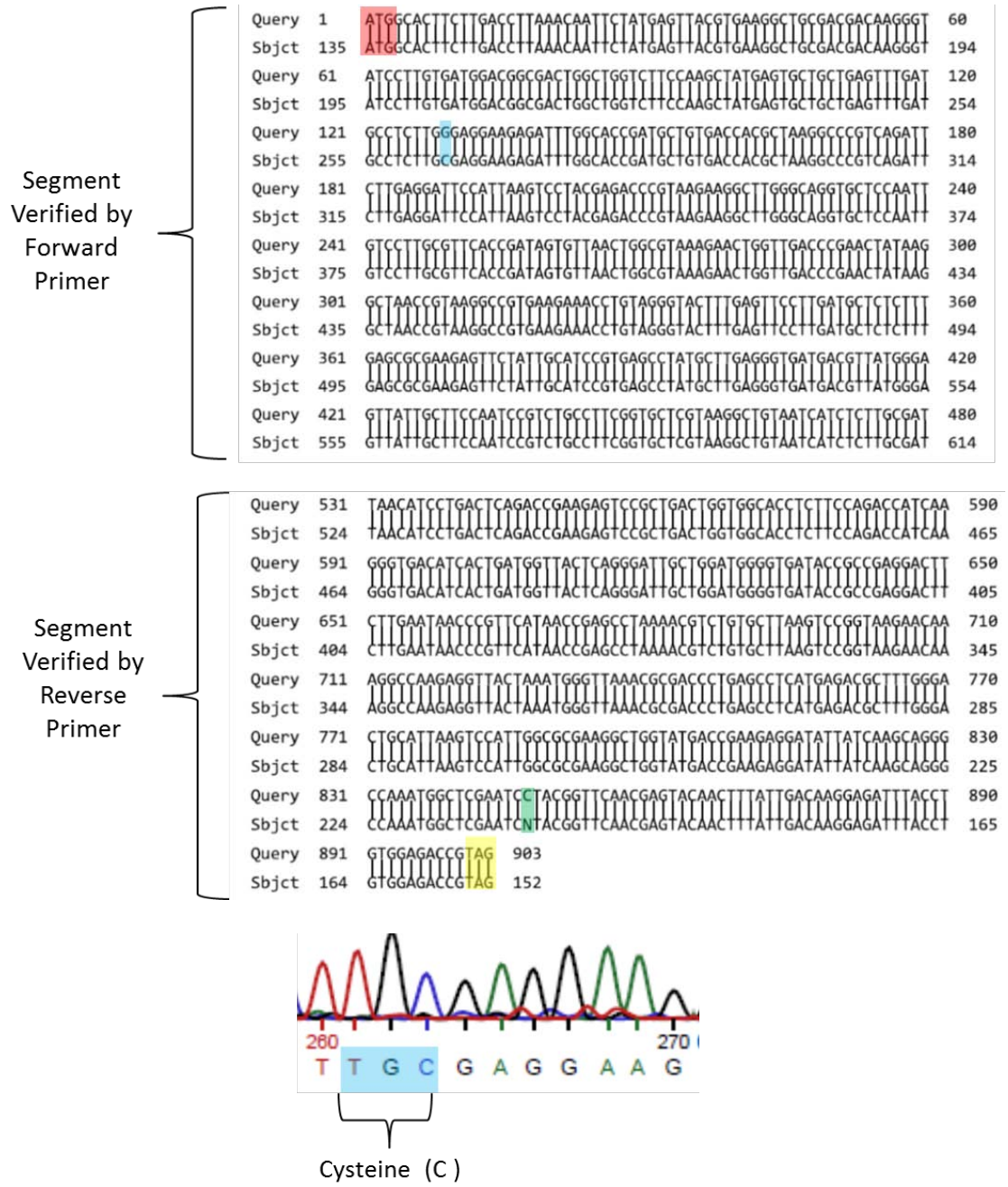

Figure 17: BLAST Results for Sample W1 and Selected Chromatogram W1 Chromatogram Segment. *This sequence analysis shows the known gene 6 sequence (Query) compared directly with the gene 6 sequence in the isolated plasmid of sample W1. The start codon is highlighted in red, and the stop codon in yellow. The mutation at base pair 129 is highlighted in light blue, this is the expected mutation that will change the codon from coding tryptophan to cysteine. The green highlighted portions are inconclusive sequence fragments from the reverse sequence that were successfully verified using the sequence resulting from the forward primer. The chromatogram segment shows the actual sequence results for the section of gene 6 that had the expected W→C mutation.*

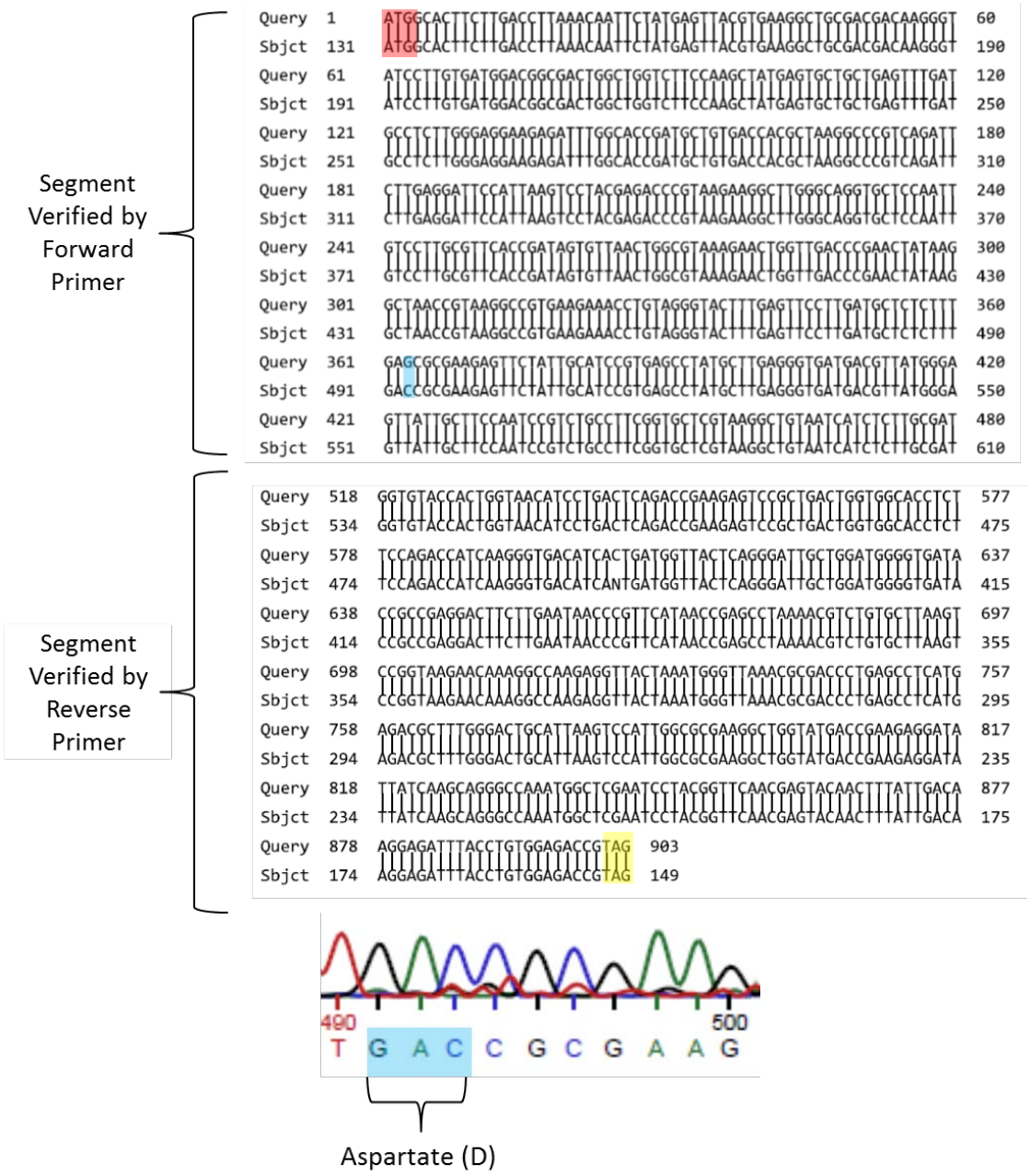

Figure 18: Sample D1 BLAST Results and Selected D1 Chromatogram Segment. *This alignment shows the known sequence of gene 6 (Query) against the sequence from the plasmid sample D1 (Sbjct). This figure shows the start codon highlighted in red, and the stop codon in yellow. The mutation at base pair 363 that will cause the codon to code for aspartate instead of glutamate is highlighted in light blue. This chromatogram segment successfully shows the codon that will now encode for aspartate, in other words, the expected E→D mutation.*

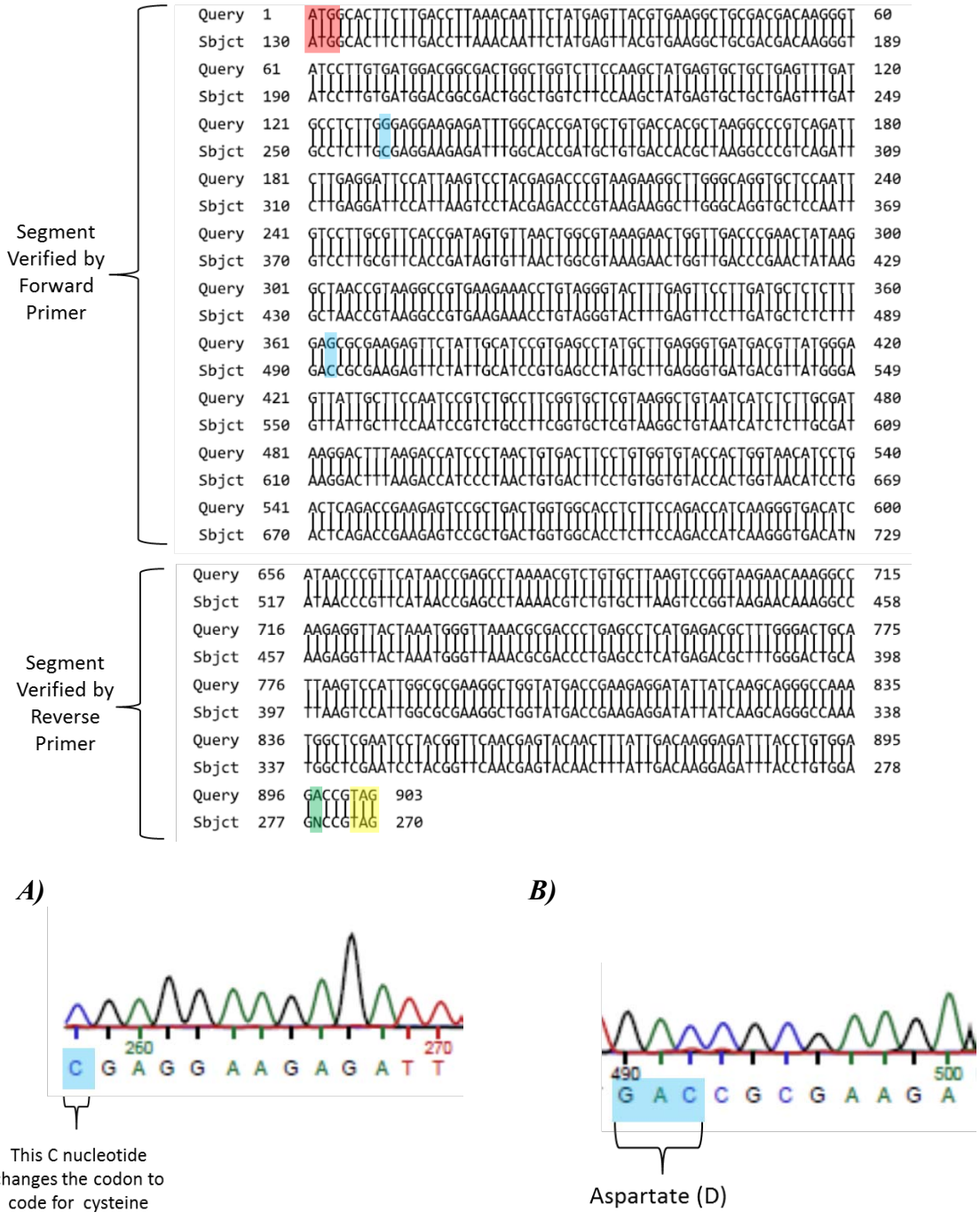

Figure 19: BLAST Analysis of Double Mutation and Select Chromatogram Segments from Sample 10. *The Start codon is highlighted in red, the two expected mutations are highlighted in light blue, the stop codon is in yellow, and a mismatched "N" highlighted in green. In the chromatogram segments, A) shows the expected W→C mutation and B) Shows the expected E→D mutation. Chromatogram results were also used to verify the correct base is present where there is an "N" at location 278 on the subject (data not shown).*

Discussion

The overall goal of this research was to create and verify plasmids that have three versions of T7 gene 6. Two of these versions should each have one point mutation, and the third should contain both of the other mutations. Throughout the creation of these plasmids, steps have been taken to ensure the procedures were successful. Gels were run to verify DNA fragment sizes after PCR I and PCR II, and to verify the amplification of the gene after transformation. Kanamycin plates were used to verify the successful transformation of the plasmids into competent cells. Finally, sequencing of isolated plasmids was analyzed to ensure the mutations were in place and the rest of gene 6 was intact.

PCR Introducing Mutations

The DNA fragments for PCR I and PCR II were verified only by evaluating their size with gel electrophoresis. This enabled the verification that the primers were likely binding at the intended sites, and that the PCR program was producing an appropriate product (Figure 12). Bands could be evaluated qualitatively for yield, by the brightness of the resulting bands. Because bands other than the expected gene fragments appear in gels for PCR I and PCR II for inserting a single mutation, it is known that other DNA is amplified by the PCR program. In PCR I, The whole T7 phage genome was used as a template, so the extra bands of DNA, all of which are larger than the gene fragments, could have been amplified from anywhere in the genome. PCR II was determined to be successful with the verification that both PCR products (with different mutations) resulted in the same sized DNA fragment that was the size of gene 6.

When introducing the second mutation onto the gene, the template used was the PCR amplified gene 6 with one mutation from the insert verification step. Using this as a template, the resulting gels were surprisingly noisy, but both PCR I and PCR II for these steps yielded the desired fragment sizes, among the others. One possible explanation for larger DNA fragments to appear would be the annealing and extension of primers to themselves. This does not usually take place when a template is available, but there should not have been much large DNA for PCR to amplify, because the template used had been isolated from agarose at the 1000 base pair mark.

Cloning Reactions and Transformation

The success of the cloning reactions and transformation was measured by the formation of colonies on kanamycin containing media. In the first round (inserting the genes that had one version of the mutation each), one plate with 50 μL and one plate with 200 μL of each of the transformations was made and incubated. It was expected that the number colonies grown would reflect the 1:4 ratio of the volumes plated. After the colonies were counted, however, the actual ratios were as follows: Tube 1 was ~1:3, Tube 2 was ~1:2, Tube 3 was ~1:2, and Tube 4 was ~1:2.5. All reactions yielded some colonies that could be utilized to create a culture.

The second round of transformation was done with plasmids containing gene 6 with both mutations. Each of the four vials of transformed competent cells were split into 50 μL and 200 μL volumes, plated, and incubated. Because of the 1:4 ratio of the volumes plated, a similar ratio was expected for the numbers of colonies that grew on the plates. The ratios observed were ~1:7 for Tube 1a, ~1:3 for Tube 1b, and ~1:6 for Tube 2a. The 50 μL plate for Tube 2b yielded no colonies, with 15 colonies on the 200 μL

plate. All reactions in round 2 yielded colonies for use in inoculating a culture, however, the second round yielded overall fewer colonies than the first round.

The difference in number of colonies between the first and second rounds of transformation could have been a result of several factors, but it is likely in any case that more competent cells took up plasmids in the first round, or that more genes were successfully inserted into plasmids in this round. Successful gene insertion may have been a function of the ideal ratios of the components of the cloning reaction. Therefore, if the PCR yield used in the reactions had a vastly different concentration from round one to round two, this could explain the difference in success. The final count of viable colonies did not impact the ultimate goal of collecting colonies to culture for plasmid purification, however, because some colonies grew for all reactions of each round, and only one colony was needed to begin a culture.

Gene Insert Verification

Before selecting plasmid samples to be sent out for sequencing, a PCR was run with the primers designed to amplify gene 6, using plasmids isolated from all six cultures with single mutations as templates. For the single mutation gene amplification, the gene was determined to be present in all samples. However, there were also very large DNA fragments (with high molecular weight), uniformly seen in all 6 lanes. These were determined to likely be the plasmid template, of which there was too much present.

The same amplification step was performed for all ten plasmid samples of the gene containing the double mutation. This gel was less uniform, but large pieces of DNA were also seen in some lanes. Lanes that only had gene 6 sized fragments (about 1000 base pairs), without any other noise, were considered ideal for sequencing. Samples were

sent for sequencing from other lanes only if they had strong evidence for the presence of the gene. Large DNA fragments appearing in this gel may also have been plasmids, present in too high a concentration. This amplification step is critical for maximizing the possibility of getting sequencing results with the desired insert, rather than wasting resources sequencing all samples, or randomly selecting samples for sequencing that may not contain the gene.

Sequencing Results

For the introduction of single mutations to the gene, plasmids were isolated for sequencing from two cultures for each of the two expected mutations, for a total of 4 sequences. These sequences were high-quality and allowed the determination that one sample of each desired mutation contained an intact gene with the mutation in the correct locus.

Seven sequences were received for the verification of the double mutation in gene 6. These sequences were of a much lower quality, and most did not contain enough full sequence to determine if the gene was intact, and the necessary mutations were present. These low quality results had many nucleotides that were unidentified, "N" or not sequenced at all. According to the sequencing party, Eurofins Genomics, possible reasons for this outcome include the DNA concentration being too low for sequencing, or the presence of contaminants such as buffer salts. The two samples that the whole gene was visible in were 4 and 10. In the BLAST of sample 4 compared with the wild-type gene 6, it was revealed that there were many mistakes in the gene, rendering it useless for the intended purposes. Sample 10, however, was determined to have a start and stop codon, both of the intended mutations at the correct loci, and all other base pairs were correct.

Plasmid sample 10 is therefore viable for later use in assays investigating the double mutation. With the information available, there is no way of determining if the other five samples that were sequenced have all the characteristics desired for future assays.

Future Directions and Conclusion

The confirmation of gp6 with the expected mutations was the most significant end result of this work. Moving forward, these plasmids can be used in complementation assays to ensure the gp6 is viable, and suppressor assays to determine its interactions with gp2.5. Complementation assays can be accomplished through infecting of *E.coli* containing plasmids with one version of mutated gp6, and using T7 that is missing gene 6. If the gp6 in the plasmids is viable with the other wild-type replisome proteins of T7, the phage will be able to replicate and form plaques.

Similarly, suppressor assays can be done by testing viability of the phage with the mutant gp2.5 and the mutant versions of gp6. This could be accomplished using E. coli that have a plasmid with the mutant gene 2.5, and one version of the mutant gene 6. These cells would be infected with T7 phages missing gene 2.5 and gene 6. If the phage was viable and able to replicate, then the gene 6 version used will sufficiently suppress the lethal mutant gp2.5 phenotype. Each of the three gp6 versions would be used separately, to determine if one of the two mutations was necessary, or both are necessary to suppress the lethal phenotype. The results of these assays will give insight to how and where these two proteins interact with each other. Based on the difference in chemical properties of the amino acids in the W42C mutation, it is predicted that it is the change that causes the suppressor phenotype. The chemical properties of the two mutations in

E120D are less different, and therefore less likely to be causing a change in how gp6 interacts with gp2.5.

Bacteriophage T7 gp6 has potential for applications in biotechnology, and is already available for purchase and use. This makes it all the more important to understand how the protein normally works, and its interactions with other proteins, to enable us to provide the best conditions for the protein to function as it is needed to in biotechnology applications.

Even for the best-characterized groups of phages, we have plenty to learn about their specific molecular mechanisms. Modern developments have shown great strides in what we know about phages in their natural roles in the environment as well as shown promising new applications, such as the utilization of phage enzymes in biotechnology protocols, their applications to combat bacteria, or the specific proposed uses of phage display to find molecules with high affinity for use as antibodies in medical applications. Phage biology is in an exciting upswing that will lead to a better understanding of global ecology and help us develop more tools from this new knowledge. The creation of the plasmids described herein is one small step toward better understanding the inner workings of phage proteins, and will allow us to have a better understanding and improved utilization of these entities.

References

Abremski, K and Hoess,R "Bacteriophage P1 site-specific recombination. Purification and properties of the Cre recombinase protein." Journal of Biological Chemistry (1984) 259: 1509-1514.

Ackermann, H. "Bacteriophage taxonomy." Microbiology Australia 32.2 (2011): 90-94

Breitburt, M., Abedon S."Phage Ecology and Bacterial Pathogenesis" Phages: Their Role in Bacterial Pathogenesis and Biotechnology. Washington, D.C.: ASM, 2005. Print.

Cai, L. "Characterization of bacteriophage T3 DNA ligase." Journal of biochemistry 135.3 (2004): 397-403.

Chibani-Chennoufi, S, "Phage-host interaction: an ecological perspective." Journal of bacteriology 186.12 (2004): 3677-3686.

Clokie, M. R., "Phages in nature." Bacteriophage 1.1 (2011): 31-45.

Cuervo, A., Pulido-Cid, M., Chagoyen, M., Arranz, R., González-García, V. A., Garcia-Doval, C. Carrascosa, J. L. Ana, "Structural characterization of T7 tail machinery reveals a conserved tubular structure among other *Podoviridae* family members and suggests a common mechanism for DNA delivery" (2013) Bacteriophage, Article Addendum.

Endy, D, Natalie K, Neal L, Stachowiak, A, Belcher, A and Banuazizi A. 20.109 Laboratory Fundamentals in Biological Engineering, Fall 2007. (MIT OpenCourseWare: Massachusetts Institute of Technology), (Accessed 10 Dec, 2014). License: Creative Commons BY-NC-SA

England, T E, Gumport R I, and Uhlenbeck O C . "Dinucleoside Pyrophosphate Are Substrates for T4-Induced RNA Ligase." Proceedings of the National Academy of Sciences of the United States of America 74.11 (1977): 4839–4842. Print.

Engler, M.J. and Richardson, C.C. (1982). P.D. Boyer(Ed.), 5, 3. San Diego: Academic Press.

Grippo, P, and Richardson C. C. "Deoxyribonucleic acid polymerase of bacteriophage T7." Journal of Biological Chemistry 246.22 (1971): 6867-6873.

Gupta, A, Oppenheim, Amos B., Chaudhary, V. K, "Phage Display: A Molecular Fashion Show" Phages: Their Role in Bacterial Pathogenesis and Biotechnology. 2005. Print.

Hamdan, S. M., and Richardson C. C, "Motors, switches, and contacts in the replisome." Annual review of biochemistry 78 (2009): 205-243.

Haq, I. U, "Bacteriophages and their implications on future biotechnology: a review." Virol J 9.9 (2012).

Hoogenboom, H.R., et al. "Antibody phage display technology and its applications." Immunotechnology 4.1 (1998): 1-20.

Hudson, JA, "Bacteriophages and the control of Bacteria in Food" Contemporary Trends in Bacteriophage Research. Contemporary Trends in Bacteriophage Research. New York, NOVA Scientific 2009. Print.

Kaźmierczak, Z., Piotrowicz, A., Owczarek, B., Hodyra, K., Miernikiewicz, P., Lecion, D., Dąbrowska, K. "Molecular imaging of T4 phage in mammalian tissues and cells." Bacteriophage 4.1 (2014): e28364.

Lee, Seung-Joo, and Charles C. Richardson. "Choreography of bacteriophage T7 DNA replication." Current opinion in chemical biology 15.5 (2011): 580-586.

Letarov, A., and E. Kulikov. "The bacteriophages in human-and animal body-associated microbial communities." Journal of applied microbiology 107.1 (2009): 1-13.

Marintcheva, B., Hamdan, S. M., Lee, S. J., Richardson, C. C. "Essential residues the C terminus of the bacteriophage T7 gene 2.5 single-stranded DNA-binding protein." Journal of Biological Chemistry 281.35 (2006): 25831-25840.

Marintcheva, B., Marintchev, A., Wagner, G., Richardson, C. C. "Acidic C-terminal tail of the ssDNA-binding protein of bacteriophage T7 and ssDNA compete for the same binding surface." Proceedings of the National Academy of Sciences 105.6 (2008): 1855-1860.

Marintcheva, B., Qimron,U., Yu, Y., Tabor, S., Richardson, C .C. "Mutations in the gene 5 DNA polymerase of bacteriophage T7 suppress the dominant lethal phenotype of gene 2.5 ssDNA binding protein lacking the C-terminal phenylalanine." Molecular microbiology 72.4 (2009): 869-880.

Matsububara, T., "Use of Bacteriophages is Display Technology: An Effective Identification of Target-Binding Peptide/Protein From Large Repertoires of Sequence." Contemporary Trends in Bacteriophage Research. New York, NOVA Scientific 2009. Print.

Örmälä A-M, Jalasvuori M. Phage therapy: Should bacterial resistance to phages be a concern, even in the long run? Bacteriophage 2013;3(1):e24219. doi:10.4161/bact.24219.

Samoylova, T, "Phage Display in Brain Tumor Research" Contemporary Trends in Bacteriophage Research. New York, NOVA Scientific, 2009. Print.

Shors, T, Understanding Viruses. 2nd ed. Sudbury, Mass.: Jones and Bartlett, 2013.Print.

Swiss Institute of BioInformatics (SIB), ViralZone, 2009.

Tabor, S. and Struhl, K. (1989). DNA-Dependent DNA Polymerases. In F. M. Ausebel, R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith and K. Struhl (Ed.), Current Protocols in Molecular Biology. 3.5.10-3.5.12. New York: John Wiley & Sons, Inc.

Turner, M. "Phage on Rampage" Nature, 2011. (Web Access Only).

Acknowledgements

I would like to extend sincere thanks to those who have contributed to the creation of this thesis. The research would not have been possible without funding through semester grants from the Adrian Tinsley Program for Undergraduate Research. I would also like to thank the Office of Undergraduate Research, especially Dr. Jenny Shanahan and Kathy Frederick, for their constant support and enthusiasm throughout my research career at Bridgewater State University, Amy Couto and Dr. Teresa King of the Honors Program for all of their support in the form of resources and tools for completing the honors requirements, Dr. Jenna Mendell, Dr. Merideth Krevosky, and Dr. Michael Carson for taking the time to review my thesis and attend the defense presentation, and lastly, a giant thank you to Dr. Boriana Marintcheva, for her guidance, patience, hard work, and time, without which this project would not have been possible.